

CLAIMS

1. An automatic keyword extraction apparatus comprising:

a first extraction means for performing an extraction of a keyword from title character string information of contents by using a first keyword dictionary in which a character string designating a sub-genre is registered; and

a second extraction means for performing an extraction of a keyword from detailed character string information of said contents by using a second keyword dictionary in which names of persons are registered and for performing an extraction of a keyword by utilizing character type separation method.

2. An automatic keyword extraction apparatus according to claim 1, wherein said first extraction means extracts a keyword from a portion within a title character string including a character string registered in a first keyword dictionary and excluding a character string registered in a predetermined character string dictionary for exclusion.

3. An automatic keyword extraction apparatus according to claim 1, wherein said first extraction means extracts as a keyword a character string separated by a special character other than Hiragana, Katakana, Chinese character, numerical character and alphabet within a title character string which includes a character string registered in a first keyword

dictionary.

4. An automatic keyword extraction apparatus according to claim 1, wherein said second extraction means performs an extraction of a keyword from a portion excluding a character string registered in a predetermined character string dictionary for exclusion by utilizing character type separation method within a remaining portion of said detailed character string information from which the keyword is extracted by using said second keyword dictionary.

5. An automatic keyword extraction apparatus according to claim 1, wherein said second extraction means treats Katakana and alphabet as the same character type while character type separation method is utilized and at the same time, treats "・" (midpoint) as Katakana or alphabet in a case when a letter just before is Katakana or alphabet correspondingly.

6. An automatic keyword extraction apparatus according to claim 1 further comprising means for downloading said second keyword dictionary via a network, wherein said second extraction means uses said downloaded second keyword dictionary.

7. An automatic keyword extraction method comprising:
a first step for performing an extraction of a keyword from

title character string information of contents by using a first keyword dictionary in which a character string designating a sub-genre is registered; and

a second step for performing an extraction of a keyword from detailed character string information of said contents by using a second keyword dictionary in which names of persons are registered and for performing an extraction of a keyword by utilizing character type separation method.

8. An automatic keyword extraction method according to claim 7, wherein a keyword is extracted in said first step from a portion within a title character string including a character string registered in a first keyword dictionary and excluding a character string registered in a predetermined character string dictionary for exclusion.

9. An automatic keyword extraction method according to claim 7, wherein a character string separated by a special character other than Hiragana, Katakana, Chinese character, numerical character and alphabet is extracted as a keyword in said first step within a title character string which includes a character string registered in a first keyword dictionary.

10. An automatic keyword extraction method according to claim 7, wherein it is performed in said second step to extract

a keyword from a portion excluding a character string registered in a predetermined character string dictionary for exclusion by utilizing character type separation method within a remaining portion of said detailed character string information from which the keyword is extracted by using said second keyword dictionary.

11. An automatic keyword extraction method according to claim 7, wherein in said second step, Katakana and alphabet are treated as the same character type while character type separation method is utilized and at the same time, " . " (midpoint) is treated as Katakana or alphabet in a case when a letter just before it is Katakana or alphabet correspondingly.

12. An automatic keyword extraction method according to claim 7, wherein further comprising a step for downloading said second keyword dictionary via a network, wherein said downloaded second keyword dictionary is used in said second step.

13. In a program of an automatic keyword extraction apparatus, a recording medium recorded with a program which can be read by a computer wherein said program comprises:

first extracting step for performing an extraction of a keyword from title character string information of contents by using a first keyword dictionary in which a character string designating a sub-genre is registered; and

a second extracting step for performing an extraction of a keyword from detailed character string information of said contents by using a second keyword dictionary in which names of persons are registered and for performing an extraction of a keyword by utilizing character type separation method.

14. A program making a computer which controls an automatic keyword extraction apparatus execute:

a first extracting step for performing an extraction of a keyword from title character string information of contents by using a first keyword dictionary in which a character string designating a sub-genre is registered; and

a second extracting step for performing an extraction of a keyword from detailed character string information of said contents by using a second keyword dictionary in which names of persons are registered and for performing an extraction of a keyword by utilizing character type separation method.